

University of Groningen

Aneuploidy in the human brain and cancer

van den Bos, Hilda

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version

Publisher's PDF, also known as Version of record

Publication date:

2017

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

van den Bos, H. (2017). *Aneuploidy in the human brain and cancer: Studying heterogeneity using single-cell sequencing*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Chapter 4

Single-cell sequencing to quantify genomic integrity in cancer

Hilda van den Bos, Bjorn Bakker, Diana CJ Spierings, Peter M Lansdorp, Floris Foijer

The International Journal of Biochemistry & Cell Biology, in press

Abstract

The use of single-cell DNA sequencing (sc-seq) techniques for the diagnosis, prognosis and treatment of cancer is a rapidly developing field. Sc-seq research is gaining momentum by decreased sequencing costs and continuous improvements in techniques. In this review, we provide an overview of recent advancements in the field of sc-seq in cancer and we discuss how sc-seq can contribute to improved care for cancer patients. Sc-seq has made it possible to study the genomes of individual cancer cells from primary tumors, metastases and circulating tumor cells, revealing inter- and intra-tumor heterogeneity, which cannot be detected using other methods. We review studies on individual human cancer cells in relation to prognosis and treatment response. Finally, future perspectives of sc-seq in cancer diagnosis and treatment are discussed with a focus on the use of circulating tumor cells to monitor therapy response and the development of personalized treatments based on knowledge about the genomic heterogeneity.

Keywords: single-cell sequencing, intra-tumor heterogeneity, tumor evolution, circulating tumor cells

Highlights

- Single-cell sequencing allows in depth analysis of the genomes of single tumor cells, which is impossible using other techniques
- Single-cell sequencing allows analysis of rare cell types such as circulating tumor cells
- Single-cell sequencing may provide future applications in the diagnosis, monitoring and treatment of cancer

Introduction

Classical sequencing of DNA extracted from millions of tumor cells provides useful information about point mutations and copy number alterations (CNAs) that are present in most tumor cells. Although tumor heterogeneity can be identified to a certain extent by sequencing multiple regions of a tumor ¹, mutations and CNAs present at low frequencies cannot be reliably identified by bulk sequencing. In contrast, sequencing of single cells (sc-seq) allows in-depth analysis of the genomes of normal and tumor cells. As the cost of sequencing the genomes of individual cells continues to decrease and the protocols become more easily accessible, the range of applications increases and the potential of sc-seq is starting to be uncovered. Such applications include quantification of aneuploidy, smaller CNA, and the mutational landscapes and thereby the genomic heterogeneity of a tumor. With this information, phylogenetic trees can be built to unravel tumor evolution. Also, rare cells isolated from biopsies or circulating tumor cells can be analyzed.

In this review, we discuss how single-cell DNA sequencing can contribute to the diagnosis, treatment and prognosis of cancer patients. Over the last several years many single-cell DNA sequencing protocols have been developed. The majority of these protocols depend on whole genome amplification (WGA) before library construction. Several WGA methods are available, each having its own advantages and disadvantages (reviewed in ²). Although WGA is necessary to generate sufficient genomic coverage for analysis of point mutations, it comes at the cost of introducing PCR amplification biases, which can obscure the detection of CNAs. To circumvent this, protocols have been developed that do not require WGA. These protocols eliminate the amplification bias, but only allow for shallow sequencing and are therefore suitable for detection of CNAs, but not for mutation analysis ^{3,4}. A major advantage of these pre-amplification-free methods is the much lower price per cell and higher throughput, allowing for many more cells to be analyzed at once. Finally, a recently published protocol allows single cell CNA analysis on formalin fixed, paraffin embedded (FFPE) tissue ⁵. Since it is common practice to store FFPE samples from tumors, this protocol opens up the opportunity to analyze an extensive repository of samples. Besides single-cell DNA sequencing, other sequencing methods that can advance cancer research have been developed, such as single-cell RNA sequencing, droplet PCR and sequencing of circulating cell free DNA. Although these methods can greatly contribute to the improvement of cancer diagnosis and monitoring, it is beyond of the scope of this review to discuss these in detail.

4

Single cell aneuploidy, CNA, and mutation analysis

Since the first report on whole genome sequencing of single breast cancer cells ⁶, a series of single cell sequencing studies have been published analyzing many different cancer types. A summary of these studies is given in Table 1. Sc-seq can reveal aneuploidy, CNAs and/or mutations in individual cancer cells from primary tumors, metastases and other (rare) tumor cell sources, such as circulating tumor cells (CTCs) and malignant pleural effusion. The heterogeneity of a tumor and/or its metastasis and CTCs can be determined, and the presence of one or more clones and evolutionary history can be mapped.

Sc-seq does not only allow the CNA pattern in a tumor or metastasis to be revealed, but also allows studies on the timing or order in which CNAs arose and subclones emerged. By sampling a tumor, its metastasis or CTC's multiple times, the response to treatment can be monitored

over time (Figure 1). Several studies have investigated the clonal evolution of tumor cells by building phylogenetic trees from this sc-seq information ^{6–11}. For example, the group of Navin has shown that breast cancer tumors experience short burst of aneuploid rearrangements followed by stable clonal expansion ^{6,10,12}. Using single-cell whole-exome sequencing on two colorectal cancer (CRC) patients, Wu et al ¹¹ demonstrated that although these CRC tumors were monoclonal in origin, several subclones were found based on accumulation of novel driver mutations.

With the use of single-cell DNA sequencing questions about the mutation frequency in tumors can be answered. Even though tumor cells typically have a higher mutation frequency than somatic cells ^{13,14}, it has been difficult to distinguish whether this is caused by an increased mutation rate or a higher proliferation rate of tumor cells as the proliferation rate is difficult to measure. While bulk sequencing has suggested a 210-fold increase in mutation rate in tumor cells compared to somatic cells ¹⁵, sc-seq suggests this rate to be much lower: at 1–13.3 times the normal mutation rate ^{10,16}. However, since only a few patients and cancer types have been investigated so far, more studies are needed to draw more general conclusions.

Circulating and disseminated tumor cells

One of the strengths of sc-seq is the possibility to analyze rare cells, such as cells isolated from a small tumor biopsy or circulating tumor cells (CTCs). The analysis of CTCs is a very promising development in the monitoring of cancer. It provides a minimal invasive way to monitor the progression of a tumor as well as the response to treatment. CTCs are isolated from the blood using an antibody staining for EpCAM and CD45 ³¹, and/or by size ³². The number of CTCs per ml of blood already has prognostic value: high numbers of CTCs are associated with poor prognosis in several cancer types, e.g. breast, prostate, lung and colorectal cancer (reviewed in ³³). Subsequent sequencing of CTCs and comparing them to primary and metastatic tumor cells yields further valuable information. Questions such as, ‘Do the CTCs have a specific karyotype/mutation landscape that allows these cells to disseminate’ or ‘Are cells at random released from the tumor?’ can be answered. Heitzer et al. ³⁴ showed that CTCs in metastatic colon cancer have similar CNA and mutation profiles. Other studies confirm similarity between the CNA profile of CTCs and the primary tumor in lung ²³ and prostate ²⁸ cancer. In contrast, Gao et al. ⁸ found the CNA pattern in CTCs to be more homogeneous than in the primary tumor of colon, breast, gastric and prostate cancer. In addition, they found the CNA pattern to resemble the pattern found in metastases, suggesting that only a certain CNA pattern allows tumor cells to be released from the primary tumor and initiate the growth of metastases. Demeulemeester *et al.* ⁹ sequenced disseminated single tumor cells (DTCs), isolated from the bone marrow of 6 breast cancer patients. Based on CNA analysis and matching phylogenetic trees the authors concluded that these DTCs separated from the primary tumor relatively late.

Table 1. Summary studies using single cell sequencing in cancer

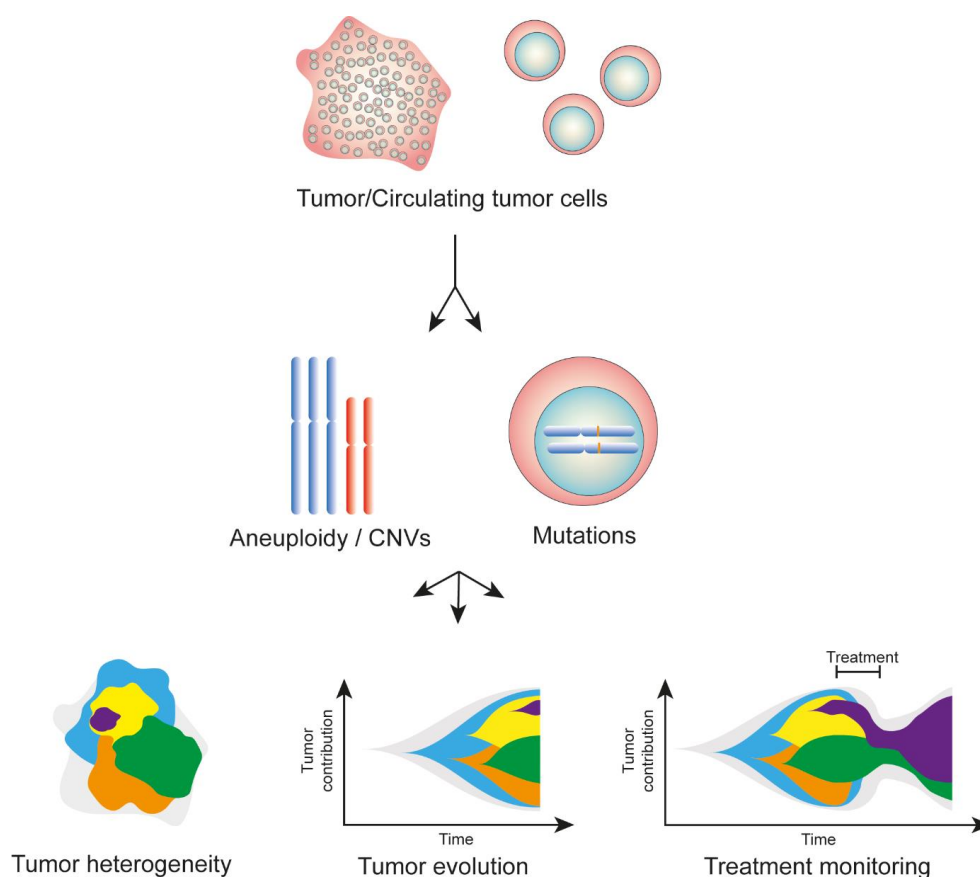
Cancer type	Tissue	# patients	# single cells sequenced, before/ after QC ¹	WGA method	Mutations /CNAs	Region sequenced	Reference
Acute lymphoblastic leukemia	Bone marrow	6	1479 cells total / ~50% passing QC	MDA (Genome Phiv2)	Mutations	Targeted sequencing	¹⁷
Acute lymphoblastic leukemia	Bone marrow	3	288 cells total/ 214 passing QC	none	CNAs	Whole genome	¹⁸
Acute myeloid leukemia	Bone marrow	3	36 cells total/ 35 passing QC	LA-PCR (PicoPLEX)	Mutations	Targeted sequencing: >1,900 loci	¹⁹
Bladder carcinoma	Primary tumor	1	66 cells total / 44 passing QC	MDA (REPLI-g)	Mutations	Exome	²⁰
Breast cancer	Primary tumor and liver metastasis	2	200 cells total/ QC NA	DOP-PCR (Genome Plex)	CNAs	Whole genome	⁶
Breast cancer	Primary tumor	2	195 cells total/ QC NA	MDA (REPLI-g)	Mutations and CNAs	Exome and whole genome	¹⁰
Breast cancer	Primary tumor	2	384 cells total/ 332 passing QC	DOP-PCR (Seqplex)	CNAs	Whole genome	²¹
Breast cancer	DTCs in bone marrow	6	63 cells total / 45 passing QC	DOP-PCR (Genome Plex)	Mutations and CNAs	Whole genome	⁹
Breast cancer	Primary tumor	12	1119 cells total/ QC NA	DOP-PCR (Genome Plex)	CNAs	Whole genome	¹²
Breast cancer	Primary tumor	4	676 cells total/ QC NA	DOP-PCR (Genome Plex)	CNAs	Whole genome	⁵
Colon cancer	Primary tumor	1	Total NA/ 67 passing QC	MDA (REPLI-g)	Mutations	Exome	²²

Table 1. (continued)

Cancer type	Tissue	# patients	# single cells sequenced, before/ after QC ¹	WGA method	Mutations /CNAs	Region sequenced	Reference
Colon cancer	Primary tumor	2	165 cells total/ 124 passing QC	MDA (REPLI-g)	Mutations and CNAs	Exome	11
Colon, breast, gastric and prostate cancer	CTCs and primary tumor	23	102 cells total / QC NA	MALBAC	CNAs	Exome and Whole genome	8
Lung cancer	CTCs	11	68 cells total/ QC NA	MALBAC	Mutations and CNAs	Exome	23
Lung cancer	CTCs	2	8 cells total / QC NA	LA-PCR (Ampli1)	CNAs	Whole genome	24
Lung cancer	CTCs	31	253 cells total / QC NA	LA-PCR (Ampli1)	CNAs	Whole genome	25
Lung cancer	Primary tumor and metastases	1	586 cells total/ 346 passing QC	None	CNAs	Whole genome	26
Myeloproliferative neoplasm	Bone marrow	1	90 cells total/ 58 passing QC	MDA (REPLI-g)	Mutations	Exome	27
Prostate cancer	CTCs	5	138 cells total / 42 passing QC	MDA (RepliPhi)	Mutations	Exome	28
Prostate cancer	CTCs	1	41 cells total/ QC NA	DOP-PCR (Genome Plex)	CNAs	Whole genome	29
Renal carcinoma	Primary tumor	1	25 cells total / QC NA	MDA (REPLI-g)	Mutations	Exome	30

¹Control cells from cell lines are not included

CNAs: copy number alterations, CTCs: circulating tumor cells, DOP-PCR: degenerate oligonucleotide primed PCR, DTCs: disseminated tumor cells, LA-PCR: linker adapter PCR, MALBAC: multiple annealing and looping based amplification cycles, MDA: multiple displacement amplification, NA: not available, QC: quality control, WGA: whole genome amplification



4

Figure 1. Single cell sequencing in cancer. By sequencing (rare) tumor cells or circulating tumor cells, aneuploidy, copy number variations and/or mutations can be identified. This enables characterization of the heterogeneity of a tumor, identification of its evolutionary path, and allows for monitoring of treatments.

Bone marrow DTCs can form a dormant reservoir that evades therapy, and might cause metastasis over time ⁹. Interestingly, sc-seq revealed that only ~50% of the cells that were previously identified as tumor cells based on morphological features, were actual tumor cells. Moreover, a number of studies have found that CTCs from certain cancer types have specific CNA patterns, which might contribute to diagnostics. A study comparing CTCs from patients with breast, gastric, prostate or colon cancer revealed reproducible CNA patterns among patients with the same cancer type, with the exception of breast cancer ⁸. The notion that breast cancer is a multi-subtype disease might explain the difference with the other cancer types studied. These observations fit well with recent mouse models for CIN cancer that show that chromosomal instability leads to aneuploid cancers that exhibit recurrent karyotypes that are tissue-specific ^{35,36}. Another study found similar CNA patterns in different patients with lung adeno carcinoma ²³. Therefore, CNA pattern in CTCs might reveal the type of cancer as well as shed light on the different cancer subtypes ⁸. Moreover, CTCs might be suitable for

non-invasive tumor monitoring over time. Dago et al. analyzed CTCs from a prostate cancer patient before and after treatment²⁹. Indeed, two therapy resistant clones were identified after treatment, one of which was already present before treatment. Importantly, another study revealed that CTCs can identify which patients will or will not respond to treatment as researchers were able to discriminate chemo sensitive from chemo refractory small-cell lung cancer based on the CNV patterns of CTCs²⁵.

In summary, sequencing CTCs holds great potential. However, it must be taken into account that capturing these cells has so far been difficult and needs further optimization. Furthermore, not all tumor cells might be captured, such as EpCAM negative tumor cells. Therefore, we need to further invest in methodologies to isolate and analyze CTCs, as they will become an essential tool in the future to monitor disease onset, progression and therapy response.

Conclusion and future perspectives

We have only just started to explore the great potential that sc-seq of cancer cells offers for expanding our knowledge on e.g. tumor evolution, metastatic potential, monitoring therapy response, and development of therapy resistance. Sc-seq can contribute to unravelling fundamental mechanisms of tumor formation and progression: the order in which tumor cells acquire CNAs and mutations can help to identify driver events. Sequencing of CTCs has the potential to provide a non-invasive method for early cancer diagnosis, monitoring therapy response and relapse, although CTCs are not found in all cancer patients.

The continuing decrease in sequencing costs, together with improved library preparation and robust whole genome amplification protocols, as well as the constantly improving toolboxes for bioinformatics analysis all help to fully exploit the potential of sc-seq. Together, these important developments will help to increase the number of cells analyzed per study as well as the number of patients included. Such larger studies involving more patients, studying larger numbers of cells, and sampling at multiple time points will allow us to study the evolution of cancers and the role of intra tumor heterogeneity at unprecedented resolution in the near future.

Funding

This work was supported by a Dutch Cancer Society grant (KWF grant RUG-2012-5549) to FF and PML, an Advanced ERC grant to PML and a Groningen Foundation for Paediatric Oncology (SKOG) grant to FF and PML.

References

1. Gerlinger, M. *et al.* Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat. Genet.* **46**, (2014).
2. Gawad, C., Koh, W. & Quake, S. R. Single-cell genome sequencing: current state of the science. *Nat. Rev. Genet.* **17**, 175–188 (2016).
3. van den Bos, H. *et al.* Single-cell whole genome sequencing reveals no evidence for common aneuploidy in normal and Alzheimer's disease neurons. *Genome Biol.* **17**, 116 (2016).
4. Zahn, H. *et al.* Scalable whole-genome single-cell library preparation without pre-amplification. *Nat. Methods* **In press**, 1–39 (2017).
5. Martelotto, L. G. *et al.* Whole-genome single-cell copy number profiling from formalin-fixed

- paraffin-embedded samples. *Nat. Med.* (2017). doi:10.1038/nm.4279
6. Navin, N. *et al.* Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90–94 (2011).
7. Greene, S. B. *et al.* Chromosomal Instability Estimation Based on Next Generation Sequencing and Single Cell Genome Wide Copy Number Variation Analysis. 1–17 (2016).
8. Gao, Y. *et al.* Single-cell sequencing deciphers a convergent evolution of copy number alterations from primary to circulating tumour cells. *Genome Res.* (2017). doi:10.1101/gr.216788.116
9. Demeulemeester, J. *et al.* Tracing the origin of disseminated tumor cells in breast cancer using single-cell sequencing. *Genome Biol.* **17**, 250 (2016).
10. Wang, Y. *et al.* Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155–160 (2014).
11. Wu, H. *et al.* Evolution and heterogeneity of non-hereditary colorectal cancer revealed by single-cell exome sequencing. *Oncogene* (2016). doi:10.1038/onc.2016.438
12. Gao, R. *et al.* Punctuated copy number evolution and clonal stasis in triple-negative breast cancer. *Nat. Genet.* **48**, 1–15 (2016).
13. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
14. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).
15. Bielas, J. H., Loeb, K. R., Rubin, B. P., True, L. D. & Loeb, L. A. Human cancers express a mutator phenotype. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 18238–42 (2006).
16. Zong, C., Lu, S., Chapman, A. R. & Xie, X. S. Genome-Wide Detection of Single-Nucleotide and Copy-Number Variations of a Single Human Cell. *Science (80-.)*. **338**, 1622–1626 (2012).
17. Gawad, C., Koh, W. & Quake, S. R. Dissecting the clonal origins of childhood acute lymphoblastic leukemia by single-cell genomics. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 17947–52 (2014).
18. Bakker, B. *et al.* Single-cell sequencing reveals karyotype heterogeneity in murine and human malignancies. *Genome Biol.* **17**, 115 (2016).
19. Hughes, A. E. O. *et al.* Clonal Architecture of Secondary Acute Myeloid Leukemia Defined by Single-Cell Sequencing. *PLoS Genet.* **10**, (2014).
20. Li, Y. *et al.* Single-cell sequencing analysis characterizes common and cell-lineage-specific mutations in a muscle-invasive bladder cancer. *Gigascience* **1**, 12 (2012).
21. Baslan, T. *et al.* Optimizing sparse sequencing of single cells for highly multiplex copy number profiling. *Genome Res.* **25**, 714–24 (2015).
22. Yu, C. *et al.* Discovery of biclonal origin and a novel oncogene SLC12A5 in colon cancer by single-cell sequencing. *Cell Res.* **24**, 701–12 (2014).
23. Ni, X. *et al.* Reproducible copy number variation patterns among single circulating tumor cells of lung cancer patients. *Pnas* **110**, 21083–8 (2013).
24. Hodgkinson, C. L. *et al.* Tumorigenicity and genetic profiling of circulating tumor cells in small-cell lung cancer. *Nat. Med.* **20**, 897–903 (2014).
25. Carter, L. *et al.* Molecular analysis of circulating tumor cells identifies distinct copy-number profiles in patients with chemosensitive and chemorefractory small-cell lung cancer. *Nat. Med.* **23**, 114–119 (2016).
26. Ferronika, P. *et al.* Copy number alterations assessed at the single-cell level revealed mono- and polyclonal seeding patterns of distant metastasis in a small cell lung cancer patient. *Ann. Oncol.* (2017). doi:10.1093/annonc/mdx182
27. Hou, Y. *et al.* Single-cell exome sequencing and monoclonal evolution of a JAK2-negative myeloproliferative neoplasm. *Cell* **148**, 873–885 (2012).
28. Lohr, J. G. *et al.* Whole-exome sequencing of circulating tumor cells provides a window into metastatic prostate cancer. *Nat Biotechnol* **32**, 479–484 (2014).
29. Dago, A. E. *et al.* Rapid phenotypic and genomic change in response to therapeutic pressure in

- prostate cancer inferred by high content analysis of single Circulating Tumor Cells. *PLoS One* **9**, e101777 (2014).
30. Xu, X. *et al.* Single-cell exome sequencing reveals single-nucleotide mutation characteristics of a kidney tumor. *Cell* **148**, 886–895 (2012).
 31. Riethdorf, S. *et al.* Detection of Circulating Tumor Cells in Peripheral Blood of Patients with Metastatic Breast Cancer: A Validation Study of the CellSearch System. *Clin. Cancer Res.* **13**, (2007).
 32. Adams, D. L. *et al.* Circulating giant macrophages as a potential biomarker of solid tumors. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 3514–9 (2014).
 33. Alix-Panabieres, C. & Pantel, K. Challenges in circulating tumour cell research. *Nat Rev Cancer* **14**, 623–631 (2014).
 34. Heitzer, E. *et al.* Complex tumor genomes inferred from single circulating tumor cells by array-CGH and next-generation sequencing. *Cancer Res.* **73**, 2965–2975 (2013).
 35. Foijer, F. *et al.* Chromosome instability induced by Mps1 and p53 mutation generates aggressive lymphomas exhibiting aneuploidy-induced stress. *Proc. Natl. Acad. Sci.* **111**, 13427–13432 (2014).
 36. Foijer, F. *et al.* Deletion of the MAD2L1 spindle assembly checkpoint gene is tolerated in mouse models of acute T-cell lymphoma and hepatocellular carcinoma. *Elife* **6**, (2017).

